# Rehabilitating Goddard: Amerindian Philology in Hupa Text Corpus Development[1]

Justin Spence
University of California, Davis

## 1.    Introduction

Michael Krauss, in his important overview of mid-20th century work on Na-Dene languages, called for the development of a "science of 'Amerindian philology'" in order to "determine much more from … older and phonologically (and otherwise also) very inadequate materials …" (1973, 923). His comments are especially relevant when considering the career of Pliny Earle Goddard, one of the most prolific scholars of Dene languages in the early 20th century. In over a quarter century of research from 1900 until his untimely death at age 59 in 1928, Goddard made important contributions to the documentation of Dene languages from all three geographic concentrations of the family. His work in areas such as morphological analysis (Golla 2003) and instrumental phonetics (McDonough and Tucker 2012) was groundbreaking, but his scholarship has not had the lasting influence of luminaries such as Sapir, Li, and Hoijer.[2] This is for a variety of reasons, one of the most important being that Goddard was, simply put, an unreliable transcriber. Anyone whose work depends on high-quality transcriptions of Dene languages must, on encountering Goddard's material for the first time, experience a sense of mild disappointment as some of his shortcomings in this regard become apparent.

Nonetheless, Goddard's scholarly contributions do have a great deal of potential to shed light on the particular languages he worked on and on comparative Dene linguistics more generally. Especially noteworthy are his collections of transcribed texts, which are among the earliest, and sometimes the only, text material available for several languages. Collected in an era when many Dene languages were still widely used in their respective communities, these collections are especially significant given current interest in adapting corpus-based methods to the study of Dene and related languages (Taff 2011, Tuttle and Lovick 2014, Crippen n.d.), and to less-studied endangered languages more generally (Ostler 2008, Mosel 2014), for linguistic research and language revitalization.  As text corpora come to play a prominent role in understanding usage-based linguistic phenomena, the painstaking effort required to develop and apply Krauss' "Amerindian philology" to interpret older documentation such as Goddard's becomes increasingly worthwhile.

The present paper describes recent efforts to do precisely this for Goddard's *Hupa Texts* (1904)*,* the earliest text collection he published, for inclusion in a searchable online dictionary and text corpus rendered in a uniform practical orthography (Hupa Online Dictionary and Texts 2008-2016). Although Goddard's transcriptions are fraught with phonological ambiguities that

[2] Krauss (1973) focused on research since 1945, but with reference to significant earlier work where it was deemed worthy of special mention. Although many of Goddard's contemporaries such as Sapir, Boas, and Morice are mentioned, Goddard is virtually invisible in his discussion.

sometimes cannot be resolved with absolute certainty, it is nonetheless possible to re-transcribe his material to arrive at plausible approximations of each word's segmental phonology, restoring contrasts that Goddard didn't account for. To date 45 of the 51 texts appearing in the collection have been processed and provisional re-transcriptions are available on the website. This work has been guided by reasonable, if not entirely unproblematic, heuristics for interpreting his transcriptions, the most important of which is the assumption of a high degree of linguistic uniformity, both within the collection and with respect to work by other researchers, unless there is a compelling evidence to the contrary. With this assumption, contrasts overlooked by Goddard can be inferred from an established understanding of Hupa grammar based on more reliable sources. Although this approach has the disadvantage of potentially obscuring real parameters of linguistic variation, it renders Goddard's materials much more usable for a variety of research and revitalization purposes than would be possible otherwise.

The paper is structured as follows. §2 provides a brief summary of texts for various languages published by Goddard during his career and discusses some general issues that are found across collections. §3 describes the Hupa Online Dictionary and Texts website, the value of incorporating material from the *Hupa Texts* into that resource, and some of the assumptions and heuristics that have guided that work. §4 considers some problematic cases where phonological ambiguities in Goddard's transcriptions cannot be resolved with certainty. §5 concludes, considering the importance of "Amerindian philology" for understanding Goddard's text collections – rehabilitating some of the deficiencies in his transcriptions, and perhaps his scholarly legacy as well.

Before proceeding, it is important to acknowledge contributions that two other researchers made towards this effort. First, Lindsey Newbold took the crucial first step of entering each of Goddard's texts into plain-text format – without her painstaking effort, none of the subsequent re-transcription would have been possible. Second, Sean O'Neill undertook a similar re-transcription effort in consultation with Hupa elders Jimmy Jackson and Calvin Carpenter in 1998-2001. O'Neill deposited his computer files with the Survey of California and Other Indian Languages and were available when the current project got underway, but technical issues with file formats and character encoding made their content inaccessible initially. It was only partway through the re-transcription process that these technical issues were solved; from that point forward O'Neill's versions became an important way of confirming re-transcriptions that the current project had produced. Thus, the versions of Goddard's texts now available on the Hupa Online Dictionary and Texts website owe a great deal to work by Newbold and O'Neill; naturally, any shortcomings are the responsibility of the present author.[3]

---

[3] The following transcription and citation conventions are used in this paper. Most examples, except those drawn from Goddard (1904), are given in the practical orthography used in the *Hupa Language Dictionary* (Golla 1996). The orthography includes English-based conventions that will be familiar to most readers, such as <ch> and <j> for postalveolar affricates, <ng> for the velar nasal, and <wh> for a voiceless labio-velar fricative. <x> and <ł> are voiceless velar and lateral fricatives, respectively. Vowel length is written with a colon <:>, but by convention is omitted in word-final position. <u> is a centralized vowel, an allophone of short /a/; the same phoneme is written <a> in some environments (e.g., before glottal stop). <'> indicates glottalization when adjacent to a tautosyllabic consonant and a glottal stop elsewhere. Following common Athabaskanist conventions, <t> and <k> are aspirated stops; <d>, <g>, and <q> are voiceless unaspirated stops. The aspirated/unaspirated contrast is neutralized in coda position; by convention, the aspirated member is written, except for the voice-valence classifier prefix *d-* in pre-stem position. <k> and <g> are front velar (written <ky> and <gy> before back vowels) and <q> is back velar

## 2.        Goddard's Text Collections

Goddard's life and professional achievements are described in detail in Kroeber (1929, 1967), Boas (1930), Dixon (1930), and Golla (2003, 2011:40-41). His interests in Dene languages started when he and his family moved to Hoopa Valley, California in 1897, where he worked as a lay missionary. He is said to have become a reasonably good speaker of Hupa while working in that capacity, and this experience prompted him to pursue graduate studies in the Department of Linguistics at the University of California, Berkeley; his was the first PhD in Linguistics awarded in North America. He remained in Berkeley for most of the first decade of the 20th century, during which time he documented Hupa and other Dene languages of the Pacific Coast region. After moving to the American Museum of Natural History in New York in 1909 and until his death in 1928, his research focus shifted to Northern and Southern Dene languages.

Over the course of his career, Goddard published collections of texts for nine Dene languages, summarized in Table 1: eleven volumes comprising approximately 1700 pages total (including translations, notes, and other explanatory matter).[4] Goddard also contributed to the publication of other texts as editor and annotator (Matthews 1907, Boas 1924), and there is a substantial quantity of unpublished material as well, especially for languages of the Pacific Coast region.[5] These collections are significant for many reasons, and Boas (1930) declared that they "will always form the basis of future studies." Especially important is the cultural information that they contain: for members of contemporary speech communities, text documentation is often much more accessible and satisfying than other kinds of material produced by professional linguists and anthropologists (decontextualized wordlists, technical analyses laden with impenetrable jargon, and the like). Moreover, as noted above, the collections published by Goddard are among the earliest text documentation for Dene languages. He was fortunate to work in a period when many of the languages were still in everyday use, even if signs of language shift were already apparent in some communities; some of the people he worked with had little knowledge of English or other European languages.[6] Insofar as Goddard's texts were transcribed before the languages had reached the state of critical endangerment (or worse) that many are in today, and when some speakers had not had prolonged exposure to European

---

(approximately uvular). See Golla (1996) for further details. Goddard's original transcriptions, when they appear in the main body of the text, are given in angle brackets (< >). When individual sounds are under consideration, they are given in square brackets ([ ]); these generally follow the conventions of the practical orthography, but in some cases for clarity glottal stop is rendered as [ʔ] and aspiration is explicitly transcribed with superscript [ʰ]. Text examples from Sapir and Golla (2001) are cited as "S&G 2001"; the text and line number are included in the reference separated by a period (so "75.3" is text 75, line 3). In examples from that source, a circumflex accent indicates "a notably higher pitch than preceding syllables" (Sapir and Golla 2001, 33), often found in contrastive contexts. Where not crucial to the discussion at hand, glosses and free translations in examples have been lightly edited for clarity.

[4] Of the varieties listed in Table 1, Hupa and Chilula are considered dialects of a single language (typically referred to collectively simply as "Hupa"), as are San Carlos and White Mountain Apache (Western Apache).

[5] Unpublished material includes Tolowa and Nongatl texts, described in Kroeber (1967) and archived at the Bancroft Library in the collection *Ethnological Documents of the Department and Museum of Anthropology, University of California, Berkeley, 1875-1958* (BANC FILM 2216). Additional unpublished texts for California Dene languages, especially dialects of Wailaki, a.k.a. "Eel River Athabaskan" (Golla 2011), are held by the American Philosophical Society. Goddard's Lassik field notebooks, which may contain texts, are part of the Melville Jacobs Papers in the University of Washington library's special collections.

[6] Cf. Spence (2016) regarding Hupa.

languages, they might be less subject to language attrition and other contact effects sometimes encountered in later documentation.

| Region | Language | Date Collected | Date Published | Page Count |
|---|---|---|---|---|
| Pacific Coast | Hupa | 1901-1902 | 1904 | 280 |
| | Kato | 1906 | 1909 | 177 |
| | Chilula | 1906 | 1914 | 93 |
| | Wailaki | 1901, 1906 | 1923 | 60 |
| Northern | Dene Sųłiné (Chipewyan) | 1911 | 1912 | 67 |
| | Tsuut'ina (Sarcee) | 1905, 1911 | 1915 | 89 |
| | Danezaa (Beaver) | 1913 | 1917 | 105 |
| Southern | Jicarilla Apache | 1909 | 1911 | 272 |
| | San Carlos Apache | 1905, 1910 | 1919 | 229 |
| | White Mountain Apache | 1910, 1914 | 1920 | 161 |
| | Navajo | 1923-1924 | 1933 | 179 |

Table 1: Summary of Text Collections Published by P.E. Goddard

Despite their rich potential, however, in practice Goddard's text collections have not figured as heavily in research on the family as work by other scholars, for several reasons. As a practical matter, Goddard's texts are rendered in an older Americanist transcription system (American Anthropological Association 1916), which makes them relatively inaccessible to anyone who is not already a specialist in early 20th century linguistic work. Re-transcribing the texts according to more familiar conventions, as the project described in the present paper seeks to do for his Hupa material, is an obvious solution to this problem, but doing so is challenging because of the more serious issue already noted above: Goddard is not a reliable transcriber of some of the fundamental phonological contrasts found in the Dene family. In his early work on Hupa and other languages from the Pacific Coast region, this is especially evident with laryngeal contrasts such as glottalized consonants and coda [ʔ] and [h].[7] In his later work outside of California and Oregon, Goddard did not transcribe contrastive tone, and he was inconsistent in his treatment of contrastive vowel length or nasalization. Goddard himself was keenly aware of his difficulties with transcription, and this is what sparked his interest in instrumental approaches to phonetics (1907, 1-2). Many of his text collections include apologies to this effect in the front matter, especially in later publications when he was more aware of the relevant contrasts but had no possibility of correcting transcriptions that had been created during fieldwork conducted years

---

[7] Goddard's inaccurate transcriptions seem to have been due, initially at least, to a rudimentary awareness of the relevant contrasts, even though many of them had already been reasonably well described in previous Americanist work (Powell 1880). In the *Hupa Texts* (1904, 95), he did include a character <*t*> (italic t) to represent "an unaspirated surd which is distinguished from *d* with difficulty." Where this letter occurs it is clearly intended to represent glottalized [t'], e.g in <*te*> for *t'e'* 'blanket' (1904, 190). Goddard transcribes this sound only rarely, however: there are many tokens where 'blanket' appears as <te> (in general, he conflates [t'] with aspirated [tʰ] rather than with unaspirated [d]). Goddard transcribes the glottalized lateral affricate consistently with <L>, noting that it is "often beginning with a slight explosion" (1904, 94). He also notes that <k>, often used in words where [q'] is expected, is "very harsh," again suggesting a nascent awareness of glottalization.

earlier.[8] Although Goddard was certainly more accurate overall than many researchers who worked in earlier decades with less training and experience, residual deficiencies in Goddard's transcriptions make materials collected in later decades, where available, more appealing sources of information for many purposes.

Another consideration is that for some languages, Goddard's worked with only a small number of speakers, sometimes only one or two. While it was not uncommon then (as now) for linguistic research on critically endangered Native American languages to be based primarily on work with just one or two speakers, Kroeber (1967, 271-272) suggests that Goddard tended to focus too narrowly on the particular details of a language as produced by a given speaker, reproducing in publications text material more or less exactly as it had been originally transcribed and translated. Kroeber notes that while this "has the virtues of a certain intensity of perception" and a vivid "savor" of a language or culture, the results are "largely unsystematic" and tend to obscure points of general theoretical or comparative interest, perhaps another reason that his work is sometimes neglected. However, in Goddard's defense, in presenting linguistic material as he had originally transcribed it, he has probably preserved phonetic and other details that another author might have omitted from a published volume. From many contemporary perspectives, Goddard's unvarnished specificity can therefore be seen as an advantage (cf. Scollon 1979a,b). This may include approaches to phonology where abstract categorical phenomena are directly linked to phonetic substance, sociolinguistic research where a "small difference of sound" (Bloomfield 1926) deemed irrelevant under mid-century phonemic theory is now grist for the theoretical mill, and current efforts to make connections between grammatical description and the data on which it is based more transparent (Berez 2015). One of the disadvantages of the approach to re-transcription taken in the project described here, where Goddard's texts are subject to a high degree of orthographic standardization, is that many of these minute details, probably in some cases reflecting heretofore unknown parameters of linguistic variation, are erased, at least in the default representations presented to users on the website.

## 3.　　Re-transcribing Goddard's *Hupa Texts*

The project described in the present paper seeks to make Goddard's *Hupa Texts* (1904), the earliest collection he published, available for linguistic research and language revitalization through the Hupa Online Dictionary and Texts website (2008-2016). Following the lead of similar online tools such as the Yurok Language Project (Garrett 2011), the overarching goal of the Hupa Online Dictionary and Texts is to compile all existing documentation for the language into a resource that can be searched through a single user interface. The text corpus component includes material from many speakers in different decades, drawn from published sources such as Golla (1984) and Sapir and Golla (2001), as well as recent texts narrated by Mrs. Verdena

---

[8] In the *Kato Texts*, Goddard acknowledges that "Dr. Edward Sapir gave assistance in some of the phonetic difficulties of this paper" (1909, 68) . In the San Carlos Apache collection (1919, 143), Goddard laments that "the recorder is called upon to hear with exactness sounds to which his ear is unaccustomed and to make distinctions which he has habitually ignored," noting nasalization of vowels and glottal stops as particular areas of difficulty. In the Wailaki Texts, originally collected in 1901 and 1906 but published only in 1923, he notes that "[t]he glottally affected sounds t' and k' are not in all places properly distinguished in these texts, due in large part to lack of care in properly entering a diacritical symbol when recording them."

Parker.[9] Tokens of words occurring in the texts are linked to a lexical database. Originally based on a learner-oriented English-to-Hupa dictionary (Golla 1996), the lexical database is organized around Hupa linguistic structure and has been expanded significantly, featuring multimedia content and more robust paradigmatic information than it was possible to compile in a print dictionary. Users can search by Hupa word or English translations, explore entries that are related to one another (in the same semantic domain, or with shared morphological structure), and move back and forth between dictionary entries and examples occurring in texts.

Given the long-term goal of providing access to all existing documentation of Hupa, Goddard's *Hupa Texts* are an important addition to this online resource. Transcribed by Goddard in 1901-1902 in Hoopa Valley, California, each text has interlinear glosses, explanatory footnotes, and a free translation occurring at the end, as in the following short example:

### XXII.

#### *Formula of the Acorn Feast.**

| ûn-Lûñ-xwed-diñ | nin-nis-an | meûk | til-tcwin-ne-en | dik-gyûñ |
|---|---|---|---|---|
| Everywhere | world | over | used to grow | here |

| na-nō-dī-ya | kyū-wiñ-ya-in-yan | dō-Lan | yū-wiñ-yûñ-hwûn-te 2 |
|---|---|---|---|
| let it come back. | People | little | must eat. |

| dō-Lan | tcū-wī-yûñ-il-he | tiñ-ye | Lan | tcū-wiñ-yûn-sil-len-te |
|---|---|---|---|---|
| Little | even if he eat | as if | much | he eats it seems. |

| kī-yauw | dō-yit-dit-te-wil-le-te | ûn-Lûñ-xoi-hwa-ûn-te | dik-gyûñ 4 |
|---|---|---|---|
| Birds | must not like it. | Every kind | here |

| ye | na-nō-dī-ya | hai | nin-nis-an | meûk | til-tcwin-ne-en. |
|---|---|---|---|---|---|
| | let it come | that | the world | on | used to grow. |

#### TRANSLATION.

#### *Formula of the Acorn Feast.*

Whatever has grown any where in the world in time past let it come back here. The people must eat but little. If one eats but little he shall feel as if he had eaten much. Birds must not like the food. Let every kind come back here which used to grow in the world.

Figure 1: Sample Text from Goddard (1904)

The longest single text in the collection is a creation story told by Emma Lewis, and the majority of the collection features genres that are familiar from ethnographic collections created in that period: traditional stories, prayers, medicine formulae, and so on. Unlike some of Goddard's

---

[9] The texts from Sapir and Golla (2001) are searchable via the website, with results presented as references to pages and line numbers containing results in the published collection.

later collections based on work with only one or two speakers, the *Hupa Texts* includes text material obtained from 14 different people (6 women and 8 men). Importantly, in 1927 Edward Sapir worked with three of the same individuals that Goddard had worked with a generation earlier (Oscar Brown, Mary Marshall, and Emma (Dusky) Frank), and some of the texts he collected (Sapir and Golla 2001) are alternative versions of ones that Goddard had transcribed. The relatively large number of speakers represented in the collection, the fact that they are virtually the earliest Hupa text material available, and the points of direct comparison with subsequent research all make them important additions to the online corpus.[10]

An important goal of the Hupa Online Dictionary and Texts website is to make it possible for users to search and retrieve texts without having to learn each of the transcription systems used in various published and unpublished resources. In Hupa, as for so many other Native American languages, there is virtually one set of transcription conventions per researcher who worked on the language. This is a general problem for people who want to work with materials that have been prepared for a given language over the course of many decades, especially new learners who may be encountering transcribed texts for the first time: it is daunting enough to navigate ways of writing an unfamiliar language in just one resource, let alone several. A sample of some differences found in three sources of Hupa documentation is provided in Table 2 (adapted from Carpenter and Spence 2015):

| IPA | tʃ | ts | l | ɬ | tɬʼ | ʔ | ŋ | ə | u | w | ʍ | tʰ | tʼ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Goddard (1904) | tc | ts | l | ʟ | ʟ |  | ñ | û | ū | w | *w*, h*w* | t | *t* |
| Sapir and Golla (2001) | *č* | *c* | *l* | *ł* | *ƛʼ* | *ʔ* | *ŋ* | *a* | *iw* | *w* | *W* | *t* | *tʼ* |
| Golla (1984/1996) | ch | ts | l | ł | tɬʼ | ʼ | ng | a/u | iw | w | wh | t | tʼ |

Table 2: Character Mismatches in Hupa Research

Notice that in addition to differences in the characters per se, readers must also determine whether or not they must attend to formatting details like italicization and capitalization. Italicization is significant in Goddard (1904), since it distinguishes [w] from [ʍ] and [tʼ] from [tʰ], but it is not significant in Sapir and Golla (2001), where the entire text line is italicized. Capitalization is significant in Goddard (1904), where it distinguishes [l] (lowercase), [ɬ] (small caps), and [tɬʼ] (capital), and in Sapir and Golla (2001), where it distinguishes [w] from [ʍ]. In Golla (1984), the first letter of each sentence is capitalized, following English spelling conventions and having no phonological significance. Navigating such dense arrays of letter correspondences is par for the course for professional linguists and experienced language teachers who must work with such materials like this out of necessity, but it is a major inconvenience for novices (whether linguists or language learners).

In order to overcome intellectual barriers that are due to different ways of writing Hupa, one of the primary challenges for incorporating texts from diverse sources into the online corpus is to convert original transcriptions to the uniform practical orthography of Golla (1996). Texts in Golla (1984) are transcribed in a precursor to the dictionary orthography, so conversion is

---

[10] Only one short text predates Goddard's material, collected by Jeremiah Curtin in 1888-1889 and archived at the National Anthropological Archives (NAA MS 2063).

straightforward (removing hyphens separating syllables, replacing short <a> with <u> except in well-defined contexts). While the transcription conventions in Sapir and Golla (2001) are quite different from the dictionary orthography, for the most part they can be converted with simple one-to-one substitutions (with some minor contextual variants).[11] To some extent this is the case for Goddard's texts as well: <ñ> can be assumed to always represent [ŋ], so it can be consistently replaced with <ng>. Goddard's <*t*> (italic t) to represent glottalized [t'] is deployed inconsistently, but where it does occur it can be replaced with <t'>. In Hupa, the aspirated alveolpalatal affricate is always labialized in syllable onsets and is transcribed by Goddard as <tcw>. Where onset <tc> does not preced <w>, it can be assumed to be glottalized and transliterated as <ch'>.[12] While attending to various contextual factors affecting transliteration is not always a trivial undertaking, in principle there are many cases where converting Goddard's transcriptions to the practical orthography can be automated (including many of the discrepancies in Table 2).

Ultimately, however, there is a significant residue of cases where Goddard did not transcribe contrasts reliably or at all, and where more interpretation is therefore needed in order to arrive at the most likely approximation of what was actually uttered by people he worked with. As noted in §2, this is the case especially for the glottalized stops [t'], [k'], and [q'] and coda laryngeals [ʔ] (<'> in the practical orthography) and [h].[13] Every time Goddard transcribes <t>, it might represent [tʰ] or [t']; where he transcribes <k>, it might represent [kʰ], [k'], or [q']; most syllables ending in a vowel might in fact be closed with [ʔ] or [h]. Because all of these sounds are exceedingly common, in practice the vast majority of words in Goddard (1904) must be inspected in order for these issues to be sorted out.

Often, an invariant word (or root) is simply listed in the dictionary, and the re-transcription can reflect the received orthographic representation when there is no evidence to the contrary. For example, Goddard's transcriptions of 'blanket' as <*te*> (with italic *t* representing glottalization) and as <te> (non-italic t representing aspiration) can both be normalized to *t'e'*, the form listed in the dictionary. However, Hupa's complex verb morphology makes it unlikely that a particular inflected form will already be listed in the dictionary. Thus, it is often necessary to apply an understanding of Hupa grammar in order to arrive at the most likely re-transcription. The customary aspect, for example, is characterized by a prefix *'i-* in the conjunct tense/aspect position; the imperfective verb stem is used, with the addition of a glottal stop if the stem ends in an open syllable, [n], or [y] (Sapir and Golla 2001:836). Goddard's <ke-it-Lō>, glossed 'she used to make baskets' (1904,189), is transparently related to imperfective *k'itł'oy* 'she weaves a basket' found in Golla (1996) – the re-transcription is *k'e'itł'o'*, with the final glottal stop

---

[11] For example, Hupa's plain vs. aspirated contrast in obstruents is neutralized in syllable codas. In Sapir and Golla (2001), the unaspirated member of each contrasting pair is written in that context; in Golla (1996), the aspirated member is. Thus <d> in the Sapir collection is replaced with <t> in coda position, except where it represents the d-classifier (retained as such presumably to indicate its etymology).

[12] In syllable onsets, Goddard transcribes the plain unaspirated postalveolar affricate as <dj>. In codas, the aspirated-unaspirated contrast is neutralized; the consonant is not labialized, and Goddard usually transcribes <tc>. In codas then, <tc> may or may not be glottalized.

[13] There are a number of other problems not considered here. For example, Goddard appears to over-transcribe vowel length, often putting a macron over vowels where a short vowel is expected, but his comments in the introduction and phonetic key to the collection suggest that the macron is sometimes intended to represent differences in vowel quality rather than length per se (1904, 92-94).

inferred from the customary aspect. Similarly, Goddard's <da-na-wil-laʟ> 'it was floating there' (1904, 325) is clearly related to the progressive theme *wila:l* 'it is floating along' (with disjunct prefixes *dah-* 'on top' and *na:-*, probably 'iterative/reversative'). The voiceless final [ł] (Goddard's small caps <ʟ>) indicates that the verb stem is in its so-called "light" form (described further below), where a long vowel *V:* becomes *Vh* (Sapir and Golla 2001, 823): the re-transcribed form is therefore *dahna:wilahł* rather than *dahna:wila:ł*. In such cases, reasonable applications of well-understood principles of Hupa grammar guide the re-transcription process.

In all of these cases, the fundamental assumption is that Goddard's failure to transcribe certain sounds are merely transcription errors, rather than reflections of real speaker variation. It is certainly possible that 'blanket' was sometimes pronounced with initial aspirated [tʰ], other times with glottalized [t'], but knowing that Goddard, by his own admission in later work, simply didn't hear this contrast well, it seems most reasonable to assume that he simply didn't write the sound as speakers produced it. That is, when confronting phonologically ambiguous transcriptions, it seems preferable to enforce orthographic uniformity rather than introducing parameters of variation that would require explanation. While this point may seem trivial, it is important to make it explicit; it is not unproblematic since it knows no a priori limitations: real parameters of previously unknown variation might be erased in this process. Put another way, where Goddard's transcriptions reflect parameters of variation that have been identified in other sources, and where there is no reason to expect that Goddard would have mis-transcribed them (e.g., cases of word-initial *ni-* vs. *'i-* discussed in Spence 2013, 90-100), we are comfortable abandoning orthographic uniformity. Other possible parameters of variation, which might become apparent only through systematic exploration of Goddard's original transcriptions, are largely excluded. The benefits of rehabilitating Goddard's transcriptions are not, therefore, without their costs.[14]

## 4.     Irreducible Ambiguity
Even with the assumption of linguistic invariance as a guiding principle of the re-transcription process, there remain a number of cases where phonological ambiguity in Goddard's transcriptions cannot be resolved with as much certainty. Two cases will be considered here: the phonological effects of so-called "light" vs. "heavy" stems (§4.1), and animate vs. inanimate subject agreement (§4.2). These are both cases where the grammatical and discourse factors governing competing plausible re-transcriptions are only partially understood. They may instead be subject to ephemeral differences in how speakers construe a given event that will be difficult to recover from information in the text. These points will be illustrated with reference to examples drawn from the highly accurate and unambiguous transcriptions in Sapir and Golla (2001).

## 4.1     Light and Heavy Stems
An important aspect of Hupa verb morphology, discussed in Golla (1977) and Sapir and Golla (2001, 823-824), is alternations involving so-called "light" versus "heavy" stems. Heavy stems indicate the diachronic presence of an enclitic short vowel *i; light stems indicate the absence of

---

[14] The text database underlying the Hupa Online Dictionary and Texts website does retain Goddard's original transcriptions, so in principle the tool could be used to facilitate such exploration, but at present this functionality has not been implemented.

this vowel. Historically, consonant-final stems lacking the "heavy" vowel underwent various phonological changes such as spirantization, shortening of long vowels, and modifications to the timing of glottal closures. The final "heavy" vowel was subsequently elided in Hupa; as a result, both light and heavy stems can occur in word-final position, and the alternation is no longer phonologically conditioned. A summary of the full range of phonological effects of the alternation can be found in Sapir and Golla (2001, 823-824). Crucially for present purposes, some of the alternations involve phonological contrasts that Goddard did not transcribe reliably. For example, long vowels in verb stems ending with *-V:wh* or *-V:s* are found as *-Vhwh* or *-Vhs* in the corresponding light form, as in (1a-b). Glottal timing is also affected, so heavy stems ending in *-V:C'* (with late glottal closure) have a light stem ending in *-V'C* (with early glottal closure), as in (2a-b):

1a. ye'inya:wh     'you come in' (imperfective)
1b. ye'inyahwh     'come in!' (imperative)

2a. me'wiłwa:tł'     'he beat against it' (perfective)
2b. me'wiłwa'tł'-te   'he's going to beat against it' (future)

In many such cases Goddard's failure to transcribe laryngeal features and segments leads to phonological ambiguity.[15] Consider the following forms transcribed by Goddard, with the stem syllable appearing in bold in each case:

3. na-tin-**dau***w*     'you better go back' (1904, 329)
4. tes-**detL**-te     'they will go' (1904, 259)
5. tce-in-**nau***w*     'she always went down' (1904, 324)

In (3)-(5), the stem syllables are ambiguous between the heavy forms *-da:wh*, *-de:tł'*, and *-na:wh* on the one hand, and the light forms *-dahwh*, *-de'tł'*, and *-nahwh* on the other. Fortunately, many of the grammatical principles governing the selection of the heavy vs. light form of the stem are well understood and easy to identify based on morphological exponents elsewhere in the verb or inferable from the English translation. Imperatives require the light stem, indicating that (3) should be re-transcribed with *-dahwh* based on the imperative force of the English gloss. Many consonant-initial enclitic elements such as the future tense *-te* also require the light stem, so (4) should be *-de'tł'*. The customary aspect, marked with *'i-* in the medial syllable in (5) and implied by "always" in the English gloss, requires the heavy stem, which can be re-transcribed as *-na:wh*.[16] In each of these examples, the heavy vs. light form can be inferred from the morphological and translational context, resolving the ambiguity (as always, with the crucial assumption that there is no variation in this regard).

---

[15] In many cases there are secondary phonological differences that help disambiguate light vs. heavy stems. For example, as noted in the example *dahna:wilahł* discussed above, heavy stems ending in *-V:l* have a corresponding light form in *-Vhł*, where devoicing of [l] makes it possible to infer the presence of the light stem (hence *Vh* rather than *V:*).

[16] The presence of the medial glottal stop can be inferred from the hyphen separating two vowels in <tce-in-nau*w*>.

However, there is a residue of cases where the factors influencing the choice between heavy vs. light stem are not well understood. Verbs in the imperfective, perfective, and progressive aspects can occur in either the heavy or the light form, with the choice "largely motivated by discourse structure" and "heavy forms of perfectives predominating in narratives" (Sapir and Golla 2001, 823). However, this tendency is not exceptionless, as the following perfective verbs with the light stem attest:

6. do: łun-ding      whił        yide:l**we'tł'**
   not many          with me     it has day-ed
   'It didn't spend many nights with me!' (S&G 2001, 453, 68.37)

7. ye:t      'e:'ng         ła'ay-xw    ninch'ing'    ya'tehs**de'tł'**
   yonder    for their part  just        down          they went
   'But people out there just kept dying off…' (S&G 2001, 312, 38.9)

The discourse factors influencing the choice of heavy vs. light stem in such cases are difficult to pin down. Golla (1977) suggests that it involves "the degree of definiteness involved in the state or activity." In some elicited examples from Verdena Parker, light forms seem to convey affective meanings such as immediacy or surprise.[17] At present It is not clear how robust this distinction is, and in any case a speaker's construal of the definiteness or immediate salience of a situation might be unlikely to have other exponents in a clause or to be reflected systematically in English glosses and free translations (the exclamation point in (6) may be an exception). These are cases where there often simply isn't enough grammatical and discourse evidence to determine whether a verb stem is in the heavy or light form, and the phonological ambiguity cannot be resolved. Given the high frequency of perfectives and the tendency for perfectives to appear with the heavy stem noted above, the present project has tended to re-transcribe them with the heavy form. Thus a perfective verb transcribed by Goddard as <xōʟ-tes-detL> 'they went' (1904, 110) is re-transcribed with the heavy stem -*de:tł'*, but absolute certainty in this and many similar cases is elusive.

## 4.2    Animate vs. Inanimate Subject Marking

Another case of ambiguity in Goddard's transcriptions that cannot be resolved with certainty involves 3rd person subject agreement. In Hupa, 3rd person subjects are coindexed on verbs according to three morphological categories, labeled "animate," "inanimate," and "obviative" in Sapir and Golla (2001, 826-827). Crucial to the present discussion is the distinction between animate, typically marked with a conjunct prefix *ch'i-*, and inanimate, which is unmarked:[18]

8a.  ch'itiliw        'he/she swims along'
8b.  tiliw            'it swims along'

9a.  ch'ichwiw        'he/she is crying'

---

[17] This is reminiscent of mirativity expressing "unexpected information" as described by Delancey (2001).

[18] The obviative subject prefix *yi-* presents significant complications to the system that are tangential to the discussion at hand since its presence generally doesn't lead to phonological ambiguity.

9b.  chwiw          'it is crying'

The animate-inanimate distinction is also reflected in agreement involving other grammatical relations: animate direct object animate *xo-*, inanimate direct object unmarked; animate postpositional object/possessor animate also *xo-*, inanimate *mi-*.

When the animate subject prefix follows an underlying long vowel, it is realized as a reduced glottal stop allomorph *'-*, in most cases as a coda on the preceding syllable, whose long vowel undergoes shortening. Compare (8a) (where the 3rd person subject prefix is word-initial) with (10) (where it follows the plural prefix *ya:-*):

10.  ya'tiliw        'they swim along'

Disjunct prefixes triggering the reduced form of the animate subject prefix, like plural *ya:-* and iterative/reversative *na:-*, are extremely common. Since Goddard does not transcribe coda glottal stops like the one in (10), this leads to a large number of verbs in the *Hupa Texts* that are phonologically ambiguous. In (11) and (12), for example, Goddard's <ya> and <na> could in principle indicate *ya'-* and *na'-* with the animate subject prefix (and concomitant shortening of the disjunct prefix vowel), or *ya:-* and *na:-* with an unmarked inanimate subject:

11.  ya-deʟ-tse       'they were living' (1904, 169)
12.  na-tes-deʟ       'they started back' (1904, 176)

Disambiguating such cases is less straightforward that one might at first suppose. Despite the labels "animate" and "inanimate," the selection of morphological category is based on a complex combination of the inherent semantic properties of a given subject and contextual factors that depend in part on a narrator's construal of a given event. Goddard himself recognized that the animate subject *ch'i-* was typically used for "adult members of the Hupa people" (1905, 99) and not with children, elderly people, non-Hupa people, and animals, essentially the analysis adopted by Golla (1970, 100). More recently, Golla adds the additional stipulation that *ch'i-* is used for "a human subject, or … a non-human subject that is the focus of the sentence or narrative" (Sapir and Golla 2001, 826).[19] By contrast, the unmarked inanimate agreement is used for instransitive subjects that are lower in animacy than a previously mentioned 3rd person subject (according to a scale involving dimensions such as living vs. non-living, human vs. non-human, plus age, gender, and foreignness among humans). The unmarked inanimate category is also often used for referents that are construed as collective plurals, and for transitive subjects that are indefinite or non-specific (Sapir and Golla 2001, 827). It has also been suggested that avoiding the animate subject prefix might sometimes be employed as a way of "deprecating" a referent whose behavior one doesn't approve of (Sapir and Golla 2001, 547, fn. 4.27). Degree of familiarity is

---

[19] "Focus" here presumably involves a high degree of discourse salience, probably akin to the definition as a "current center of attention" given in Gundel, Hedberg, and Zacharski (1993), whether or not the subject is the focus of the sentence in an information-theoretic sense.

another dimension that is probably relevant to determining whether something is treated as animate vs. inanimate.[20]

The recognition that discourse and other factors play a role in determining which morphological category is used in subject agreement is important: the selection of animate vs. inanimate does not depend entirely on the inherent semantics of the referent triggering the agreement. For example, the noun *k'iwinya'nya:n* 'Indian person' is frequently used with unmarked inanimate subject agreement, as in (13), where the verb is inanimate *na:diɫ* rather than animate *na'diɫ*:[21]

13.  … niwhong-xw    na:diɫ-te                    k'iwinya'nyâ:n
     … well           things, people will go about   Indians
     '… [he talks about] people being well.' (S&G 2001, 57, 2.18)

Such cases are analyzed as involving collective action. However, while this seems to be a very robust tendency, *k'iwinya'nya:n* is not inherently collective. In (14), the verb *ch'initindil* includes the animate subject prefix *ch'i-*, agreeing with *k'iwinya'nya:n*:

14.  hayahujit  hayah  ch'initindil        k'iwinya'nya:n  hay  ya'diwilye'-te
     Then        there   they arrive there   people, Indians  the  ones who are going to dance
     'Then the people who are going to dance arrive there.' (S&G 2001, 60, 2.50)

This form in (14) is incompatible with a collective interpretation, since it explicitly includes a distributive plural prefix *ti-* as well; the referent is also definite ('the ones who are going to dance') in this case, perhaps favoring animate rather than inanimate agreement. In (15), the verb *na'wa* is marked with the animate subject prefix; the corresponding inanimate form would be *na:wa*. Although the verb is glossed in English as plural 'are going about', the stem *-wa* is inherently singular 'one person goes' rather than plural *-dil*:[22]

15.  … k'[iwinya'nya:n]     xa'unɫu(ng)   na'wa
     … people                so many        are going about
     '… so many people [are going] around!' (S&G 2001, 374, 50.12)

The use of singular verb roots with plural subjects indicates a collective plural interpretation (cf. Sapir and Golla 2001, 568, fn. 11.44; 657, fn. 50.10), but in (15) the subject agreement is

---

[20] At the 2009 Athabaskan Languages Conference in Berkeley, Golla asked Verdena Parker, who was in the audience, whether or not someone could use a verb inflected with the animate subject prefix to talk about an activity performed by an ant. Her response, paraphrasing somewhat, was that it depends on how well you know the ant! She has made similar statements numerous times in elicitation sessions as well.

[21] This tendency is apparent even when the noun is not in its explicitly collective form *k'iwinya'nyun*, historically *k'iwinya'nya:n* + *ni*, a collectivizing plural suffix (cf. Sapir and Golla 2001, 860). It is also worth pointing out that explicitly collective nouns marked with collective *-ni* are not incompatible with the animate subject prefix *ch'i-* (or its reduced allomorph), as in examples (18) and (19), where *tsumehstɫ'on* with short [o] in the last syllable is derived from singular *tsumehstɫ'o:n* + *ni*.

[22] Square brackets in this example indicate that Sapir only transcribed *k'* here, and the rest of the word was inserted by the editors; parentheses involve the editorial restoration of a consonant that is deleted due to the following word with initial [n] (Sapir and Golla 2001:29).

nonetheless animate rather than inanimate. Thus, while a noun like *k'iwinya'nya:n* might tend toward a collective interpretation, and hence inanimate subject agreement, the collective interpretation is not required (as in (14)), nor does a collective interpretation guarantee inanimate agreement (as in (15)).

An alternative heuristic for resolving ambiguities with respect to the presence or absence of the animate subject prefix is to consider how a given referent is marked in the surrounding narrative context. For example, the first text in Goddard (1904), Emma Lewis' creation story featuring the "creator and culture hero" Yima:ntiw'winyay, opens with the sentence in (16):

16. tcō-xōʟ-tcwe-diñ          e-nañ          na-teʟ-di-tcwen
    Myth-place               it was         he grew
    'It was at Tcōxōltcwediñ he came into being.' (Goddard 1904, 96)

The verb transcribed as <na-teʟ-di-tcwen> is ambiguous: it could be animate *na'tehłdichwe:n* or inanimate *na:tehłdichwe:n.* Inspecting subsequent lines, the same referent is indexed with unambiguously animate morphology: <tcis-lan> 'he was born' with the animate subject prefix *ch'i-*, <xō-la> 'his hand' with animate possessor prefix *xo-*, etc. The immediate discourse context thus provides clues to help resolve ambiguities in Goddard's transcription.

However, this disambiguation strategy also is not without problems, since there are cases where the morphological animacy of a referent varies even within relatively local stretches of the same text. This is illustrated in (17)-(19):

17. … hayah-mił   ch'e'te:de:tł'        hay   xo'osday
    thereupon        they all went out   the    men
    '… [then] the men went out.' (S&G 2001, 505, 75.8)

18. haya:ł   hay   tsumehstł'ôn   midiłwa        ch'e:te:de:tł'
    then      the    women          in their turn   they all went out
    'Then, in their turn, the women went out.' (S&G 2001, 505, 75.10)

19. yiwiding-hit   'aht'ing   yehna'te:de:tł'          hay   tsumehstł'on
    finally           all          they all come back in   the    women
    'After a while all the women come back in.' (S&G 2001, 505, 75.12)

In (17), the verb *ch'e'te:de:tł'* 'they all went out' is marked with the animate subject prefix, agreeing with *hay xo'osday* 'the men'. Two lines later, in (18) the same verb is in the unmarked inanimate subject form *ch'e:te:de:tł'*, agreeing with *hay tsumehstł'on* 'the women'. According to an analytic footnote to the text, the use of the unmarked inanimate subject here is due to the fact that the subject 'women' stands in contrast to the men having performed the same activity immediately beforehand (Sapir and Golla 2001, 712, fn. 75.16) – an example of the inanimate agreement being used for a subject that is less animate than a previously-mentioned one. Two lines later, in (19), *tsumehstł'on* is treated as animate, where the verb *yehna'te:de:tł'* 'they all come back in' has the reduced allomorph of the animate subject prefix (instead of inanimate *yehna:te:de:tł'*).

A similar example is found in a text describing funerary practices (Sapir and Golla 2001, 190, text 20). In line 24, the text describes how a medicine woman leads a group down to the river, and a gravedigger follows her: *ma:-xoda:'unâ:wh* 'she (customarily) leads them down', with unmarked inanimate subject, followed in the next clause by *xoq'eh xoda'unâ:wh* 'following her – he (customarily) goes down', with the reduced form of the animate subject prefix apparent due to concomitant shortening of the long vowel of the prefix *xoda:-* 'downhill'. Importantly, in the second clause reference to the medicine woman that the gravedigger follows is indexed with the animate postpositional object prefix *xo-* on *xoq'eh* 'following her'. In this case, the same referent is treated as inanimate as a subject in one clause, and as an animate possessor in the very next one. This could be another example of a contrastive subject involving an animacy scale as in (17) and (18), but here the verb with the unmarked inanimate subject precedes rather than follows the one marked with the animate subject prefix.

The upshot of these examples is that neither the inherent properties of particular nominals nor the animacy status of a given referent in a local narrative context is a foolproof way to disambiguate with certainty cases like (11) and (12) above. Certainly there are strong tendencies – adult human subjects who are core protagonists in narratives are indeed typically tracked with animate agreement prefixes. But there are enough complicating factors such as contrast, relative animacy, definiteness, and collectivity that often make it difficult to establish with certainty how a narrator might have construed a referent. Currently the project is working towards examining the discourse context of each ambiguous example in an effort to make an educated guess as to the most plausible scenario; cases where this hasn't yet been done are currently indicated with square brackets in re-transcriptions, e.g., *na[']tehsde:tł'* for (12) above. Ultimately, the degree of certainty associated with particular re-transcribed forms will often be much lower than in other cases where the grammatical basis for disambiguation is clearer.

## 5.    Conclusion

Despite residual difficulties disambiguating some of Goddard's transcriptions in the *Hupa Texts*, the re-transcribed versions now available on the Hupa Online Dictionary and Texts website do make it much easier to locate information of interest than was previously possible. Certainly they must be used with some caution – one would need to think carefully before incorporating them into a corpus-based study of animate vs. inanimate agreement in 3$^{rd}$ person reference tracking, for example, or of exceptional cases of light vs. heavy stems, or of phonological variation. But for many other kinds of linguistic analysis, and especially as a resource for language revitalization, the re-transcribed *Hupa Texts* are a step in the right direction.

The re-transcription effort reported here is fortunate to have resources that don't exist for all the other languages that Goddard worked on: the previous work done by Sean O'Neill with fluent speakers as a way to confirm proposed re-transcriptions, and the possibility of checking remaining uncertainties with Mrs. Parker (although other documentation priorities limits how often we are able to work on Goddard's material together). Nonetheless, assumptions of linguistic uniformity and the use of grammatical and discourse context as a means of overcoming some of Goddard's shortcomings as a transcriber of Hupa point towards general principles of philological interpretation that will be applicable to other text collections that Goddard produced during his career. Certainly there seems to be a fair bit of interest in such an enterprise at the moment, with similar efforts underway for Wailaki (Kayla Begay and Cheryl Tuttle), Dene

Sųłiné (Sally Rice and John Janvier), and Tsuut'ina (Chris Cox and Bruce Starlight). If Krauss' "Amerindian philology" can ever be usefully developed with Pliny Earle Goddard's scholarly legacy in mind, the time is certainly ripe to do so.

References

American Anthropological Association. 1916. *Phonetic Transcription of Indian Languages: Report of the Committee of the American Anthropological Association*. Washington, DC: Smithsonian Institution.

Berez, Andrea L. 2015. "Reproducible Research in Descriptive Linguistics: Integrating Archiving and Citation into the Postgraduate Curriculum at the University of Hawai'i at Mānoa." In *Research, Records, and Responsibility*, edited by Nick Thieberger and Linda Barwick, 39-51. University of Sydney Press.

Bloomfield, Leonard. 1926. "A Set of Postulates for the Science of Language." *Language* 2(3):153-164.

Boas, Franz. 1924. "Ts'ets'aut, an Athapascan Language from Portland Canal, British Columbia." Arranged and annotated by Pliny Earle Goddard. *International Journal of American Linguistics* 3(1):1-35.

Boas, Franz. 1930. "Pliny Earle Goddard." *International Journal of American Linguistics* 6(1):1-2.

Carpenter, Kayla and Justin Spence. 2015. "Beyond Words: Enhancing a Learner's Dictionary with Texts in Hupa (Athabaskan)." Paper presented at the 2015 SSILA Winter Meeting, Portland, OR.

Crippen, James. n.d. "Text Corpus of the Tlingit Language for Linguistic Research." Online resource: https://github.com/jcrippen/tlingit-corpus

Delancey, Scott. 2001. "The Mirative and Evidentiality." *Journal of Pragmatics* 33:369-382.

Dixon, Roland B. 1930. "Pliny Earle Goddard." *Proceedings of the American Academy of Arts and Sciences* 64(12):526-528.

Garrett, Andrew. 2011. "An Online Dictionary with Texts and Pedagogical Tools: The Yurok Language Project at Berkeley." *International Journal of Lexicography* 24(4):405-419.

Goddard, Pliny Earle. 1904. *Hupa Texts*. University of California Publications in American Archaeology and Ethnology 1(2):89-368.

Goddard, Pliny Earle. 1905. *Morphology of the Hupa Language*. University of California Publications in American Archaeology and Ethnology 3.

Goddard, Pliny Earle. 1907. *Phonology of the Hupa Language, Pt. 1: The Individual Sounds*. University of California Publications in American Archaeology and Ethnology 5(1):1-19.

Goddard, Pliny Earle. 1909. *Kato Texts*. University of California Publications in American Archaeology and Ethnology 5(3):65-238.

Goddard, Pliny Earle. 1911. *Jicarilla Apache Texts*. Anthropological Papers of the American Museum of Natural History 8.

Goddard, Pliny Earle. 1914. *Chilula Texts*. University of California Publications in American Archaeology and Ethnology 10(7):289-379.

Goddard, Pliny Earle. 1915. *Sarsi Texts*. University of California Publications in American Archaeology and Ethnology 11(3):189-277.

Goddard, Pliny Earle. 1916. *Beaver Texts*. Anthropological Papers of the American Museum of Natural History 10(5):295-397.

Goddard, Pliny Earle. 1919. *San Carlos Apache Texts*. Anthropological Papers of the American Museum of Natural History 24(3):141-368.

Goddard, Pliny Earle. 1920. *White Mountain Apache Texts*. Anthropological Papers of the American Museum of Natural History 24(4):369-528.

Goddard, Pliny Earle. 1923. "Wailaki Texts." *International Journal of American Linguistics*. 2(3/4):77-135.

Goddard, Pliny Earle. 1933. *Navajo Texts*. Anthropological Papers of the American Museum of Natural History 34(1):1-179.

Golla, Victor. 1970. "Hupa Grammar." PhD diss., UC Berkeley.

Golla, Victor. 1977. "A Note on Hupa Verb Stems." *International Journal of American Linguistics* 43(4):355-358.

Golla, Victor. 1984. *Hupa Stories, Anecdotes, and Conversations*. Hoopa, CA: Hoopa Valley Tribe.

Golla, Victor. 1996. *Hupa Language Dictionary*. 2nd. ed. Hoopa, CA: Hoopa Valley Tribal Council.

Golla, Victor. 2003. "Pliny Earle Goddard, Hupa and the History of Athabascan Linguistics." In *Proceedings of the 2003 Athabaskan Languages Conference*, compiled by Siri Tuttle, 1-8. Alaska Native Language Center Working Papers 3. Fairbanks: ANLC.

Golla, Victor. 2011. *California Indian Languages*. Berkeley: University of California Press.

Gundel, Jeanette K., Nancy Hedberg, and Ron Zacharski. 1993. "Cognitive Status and the Form of Referring Expressions in Discourse." *Language* 69(2):274-307.

Hupa Online Dictionary and Texts. 2008-2016. http://nalc.ucdavis.edu/hupa/hupa-lexicon.html.

Krauss, Michael. 1973. "Na-Dene." In *Linguistics in North America*, edited by Thomas A. Sebeok, 903-978. Current Trends in Linguistics 10. The Hague: Mouton.

Kroeber, Alfred L. 1929. "Pliny Earle Goddard." *American Anthropologist* 31(1):1-8.

Kroeber, Alfred L. 1967. "Goddard's California Athabascan Texts." *International Journal of American Linguistics* 33(4):269-275.

Matthews, Washington. 1907. *Navaho Myths, Prayers, and Songs with Texts and Translations*. Edited by Pliny Earle Goddard. University of California Publications in American Archaeology and Ethnology 5(2):21-63.

McDonough, Joyce and Benjamin V. Tucker. 2012. "Replicating Goddard: A Contemporary Airflow and EGG Study of Dene Sųłiné." With Valerie Wood and Horace Adams. *University of Rochester Working Papers in the Language Sciences*. Online resource: http://www.sas.rochester.edu/lin/joycemarymcdonough/mcdonoughtucker.pdf

Mosel, Ulrike. 2014. "Corpus Linguistic and Documentary Approaches in Writing a Grammar of a Previously Undescribed Language." In *The Art and Practice of Grammar Writing*, edited by Toshihide Nakayama and Keren Rice, 135-157. Language Documentation & Conservation Special Publication 8.

Ostler, Nicholas. 2008. "Corpora of Less Studied Languages." In *Corpus Linguistics: An International Handbook, Volume 1*, edited by Anke Lüdeling and Merja Kytö, 457-483. Berlin: Mouton.

Powell, John Wesley. 1880. *Introduction to the Study of Indian Languages with Words, Phrases, and Sentences to Be Collected*. Washington, DC: Government Printing Office.

Sapir, Edward and Victor Golla. 2001. "Hupa Texts, with Notes and Lexicon." In *Collected Works of Edward Sapir, Vol. 14: Northwest California Linguistics*, edited by Victor Golla and Sean O'Neill, 19-1011. New York: Mouton de Gruyter.

Scollon, Ronald. 1979a. "Variable Data and Linguistic Convergence: Texts and Contexts in Chipewyan." *Language in Society* 8(2):223-243.

Scollon, Ronald. 1979b. "236 Years of Variability in Chipewyan Consonants." *International Journal of American Linguistics* 45(4):332-342.

Spence, Justin. 2013. "Language Change, Contact, and Koineization in Pacific Coast Athabaskan." PhD diss., UC Berkeley.

Spence, Justin. 2016. "Lexical Innovation and Variation in Hupa (Athabaskan)." *International Journal of American Linguistics* 82(1):71-93.

Taff, Alice. 2011. "Deg Xinag Xidhoy: Deg Xinag Stories." Online 4.5 hours of Deg Xinag Athabascan traditional stories, history, and personal narrative, audio with time-aligned transcription/translation. University of Alaska Southeast.

Tuttle, Siri and Olga Lovick. 2014. "Alaskan Athabascan Commands: Grammatical Documentation from a Database Project." Paper presented at the 2014 SSILA Winter Meeting, Minneapolis, MN.